US03/41426

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(54) Title: SYSTEM AND METHOD FOR MAPPING OF MULTIPLE TRAIT COMPLEXES IN MULTIPLE ENVIRONMENTS
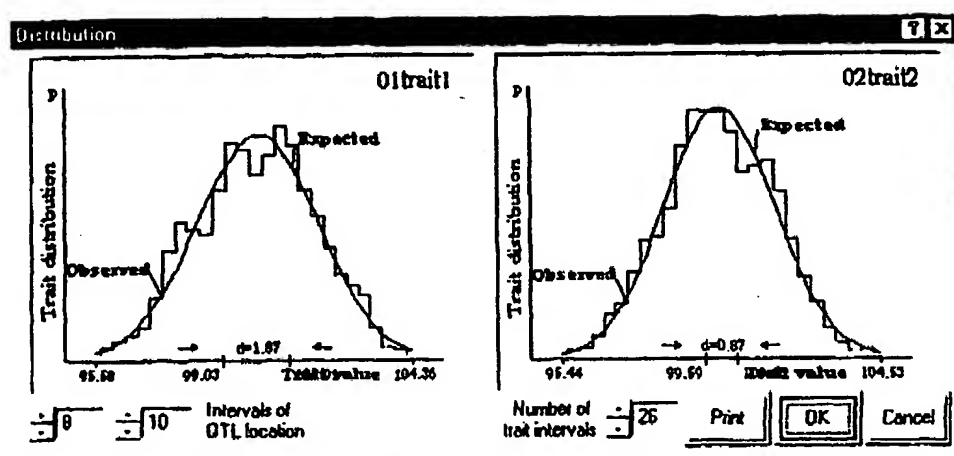
(57) Abstract: A method and system for mapping quantitative trait loci in mixed plant or animal populations for improving the efficacy of plant or animal breeding, and for effective regulation of diagnostic and therapeutic genomics in medical genetics. The method employs algorithmic models to predict the association of genetic markers with a desired phenotypic trait by taking account of the effect of variance and covariance of the analyzed QTL. These models allow the designing of new mapping frameworks and simulation tools, and the association to be extrapolated to the progeny of the plant, animal or genetic material tested as well as to multiple environments.

WO 01/88086 A2

SYSTEM AND METHOD FOR MAPPING OF MULTIPLE TRAIT COMPLEXES IN MULTIPLE ENVIRONMENTS

1. CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of U.S. Provisional Patent Application No. P -2428, filed April 20, 1999, and which is incorporated by reference in its entirety herein.

2. BACKGROUND OF THE INVENTION

The present invention relates to a system and method for genetic mapping of Trait Loci (QTL) in plants and animals. An organism's properties (traits) depend on genes and the environment. Most of the traits of economical and medical importance are Quantitative Traits. It is estimated that 98% of the economically important phenotypic traits in domesticated plants are quantitative traits. The chromosomal location or a gene or a group of closely linked genes that affects a trait (that is measured on a quantitative scale) is referred to as Quantitative Trait Locus. The quantitative traits are typically affected by more than one gene and are sensitive to the environmental conditions. Information of the chromosomal position of QTLs is of primary importance for breeding, ecology and medical genomics. For example, gene mapping provides information about the position on a chromosome of different genes thus enabling procedures to replace specific genes using marker assisted selection (MAS) or cloning by genetic engineering.

With the near completion of the mapping of the human genome, thousands of gene sequences have become available for further study in databases in both the public and private domains. There is a race to map the genes of various organisms in the pursuit of knowledge to improve agricultural products, control diseases in plants, animals and humans, and to safeguard the ecosystem.

It is estimated that 98% of the economically important phenotypic traits in domesticated plants are quantitative traits. QTL include genes that control numerically representative phenotypic traits that are usually continuously distributed within a family of individuals as well as within a population of families of individuals.

The simplest experimental paradigm developed to analyze QTL involves establishing a mapping population by controlled crosses of inbred lines, obtaining segregating progeny, genotyping multiple marker loci and evaluating one to several quantitative phenotypic traits among the segregating progeny obtained. The QTL are then identified on the basis of significant statistical associations between the genotypic values (marker scores) and the quantitative trait values (phenotypic scores) among the segregating progeny. This experimental paradigm is ideal in that the parental lines of the F1 generation have the same degree of linkage, all of the associations between the genotype and phenotype in the progeny are informative and linkage disequilibrium between the marker loci and phenotypic traits is maximized. However, because of usual limitations on the sample size of progeny studied, the paradigm described above lacks the necessary statistical power to identify QTL for most traits of economic and medical importance in breeding, ecology and medical genetics. This lack of statistical power produces biased estimates of the QTL that are identified.

New techniques of molecular marking have dramatically increased the mapping resolution. A new era of breeding, based on Marker-Assisted Selection has been started; QTL mapping is one of the central components of this new breeding technology. The list of major QTL mapping packages include: MapmakerQTL, QTL Cartographer, MapQTL, PLABQTL, and MultiCrossQTL.

CartographerQTL, is the only package that allows for multiple trait analysis, but for a very limited range of population structures. It does not allow for selective genotyping of correlated traits.

MapQTL, is the only package that takes advantage of combined treating of multiple environmental data, but only in situations where the environments can be characterized by some 'physical' attributes.

MultiCrossQTL provides the best service for treatment of different population structures, but does not include multiple trait analysis and multiple environment analysis.

None of the above packages takes into account the variance or covariance effect of the analyzed QTLs, which seriously reduces the power of accuracy. Additionally, none provides simulation assistance for the "ongoing experiments" to allow the correction of the experimental design based on the available information, at any project step, to optimize the decision making on the succeeding steps.

None of the above packages allows the inclusion into the prepared framework new mapping models and population structures invented by the user. And finally, none of the available packages provides the spectrum of analytical and simulation tools, which can be used for comparison of different mapping designs.

The vast majority of traits that are the subject of breeding efforts for both plants and animals are Quantitative Traits. These include yield and productivity and their numerous components, yield quality, the developmental characteristics, resistance to stressful conditions, pests, and diseases, the ability for efficient utilization of favorable conditions (e.g., fertilizers and water), suitability to modern food technology, etc. In the last decade, large scale genetic mapping efforts started on many dozens of agricultural plants and animals with hundreds of molecular

markers like enzyme loci, restricted fragment length polymorphisms (RFLP) and polymerase chain reaction markers (PCR), like RAPDs, ISTR, STS, and especially SSR (simple sequence repeats, or microsatellites). Especially promising is a new class of molecular genetic markers, referred to as SNPs (single nucleotide polymorphism) *(Nielsen R., 2000, Genetics, 154, 931-942)*. New techniques of molecular marking have dramatically increased the resolving power of genetic analysis, opened up unsuspected opportunities for modernization of agriculture, agrobiotechnology and agroindustry, primarily due to the breakthroughs in breeding technology. These are based on the rapid progress in understanding the genome structure and functional organization, new insights into the problems of domestication based on comparative genetics of genomes, new tools for positional cloning of genes with qualitative effect, and good prospects for genetic dissection of quantitative traits with the subsequent cloning of corresponding genes. *Paterson et al., 1988, 1990; Beckmann and Soller, 1991; Paterson A.H. et al., 1988, Nature 335:721-726; Paterson A.H. et al., 1990, Genetics 124:735-742; Beckman JS and Soller M, 1991, Theor. Appl. Genet. 74:369-378; Korol AB et al., 1994, Recombination Variability and Evolution, Chapman & Hall, London.* Consequently, a new era of breeding has already started, based on Marker Assisted Selection, that should revolutionize the field. QTL mapping is one of the major scientific components of this new breeding technology.

Many dozens of cultivated plants, domestic animals, and the major forest species are already involved in this process as well as important wild species and endangered (disappearing) organisms. No doubts, such efforts will be spread on all economically and many ecologically important species. That means a broad spectrum of qualitatively different new genetic situations arise which should be dealt

with. Likewise, new methods appear in the scientific literature that may and should be integrated in ongoing mapping projects and designs, allowing for a further elevation of the performance. The cost of the mapping experiments and the expected profit from corresponding map-based breeding efforts are by far outweighing any additional expenses of sophisticated data analysis. Hence, any complication of QTL mapping models is justified, in spite of higher CPU time requirements, if it allows for a further increase in the QTL detection power and mapping accuracy. Consequently a lot of work remains to be done in order to extract fully the mapping information hidden in the collected data, and ensuring more efficient application of this information in further breeding. New mapping methods and algorithms that increase the resolution, should urgently be converted into user-friendly and broadly available technology, clear to the breeder, in spite of its complicated genetic statistical and computational background.

Current QTL packages include: MapmakerQTL, QTL Cartographer, MapQTL, PLABQTL, and MultiCrossQTL. Of these, only CartographerQTL provides the possibility of multiple trait analysis, for a very limited range of population structures – F2, backcross and recombinant inbred lines. This package is based on the method of multiple trait interval mapping proposed by *Jiang, C. and Zeng, Z-B (Genetics 140:1111-1127 (1995))* and takes advantage of an additional increase in the mapping resolution due to regression co-factors included into the model *(Zeng, Z-B, 1994, Genetics 136:1457-1468)*. However, because of the limitations of the regression mapping approach employed in this package, it does not allow the use of selective genotyping for correlated traits, one of the most efficient ways to increase the efficiency of large-scale QTL mapping and fine mapping. The reason is that with correlated quantitative traits, selective genotyping may result in biased parameter

estimates and as a result limits its application. This is especially true in case of the variance-covariance effects of the target QTLs. However, recent studies allowed the correction of results when the maximum likelihood approach was applied. Moreover, further studies indicate that this method may be useful in situations involving pooled DNA analysis, by applying the method of sequential estimation of linkage under censored sampling *(Ronin YI et al., 1996; Biometrics 52:1428-1439).*

Beside mentioned limitations, the QTL Cartographer as well as other packages assume no variance or covariance effect of the analyzed QTLs, i.e., the residual variance (covariance) of the quantitative trait(s) is assumed to be the same in the alternative QTL groups. This may not be the case, and therefore by using an incorrect model serious error is introduced. On the other hand, by taking into account the effect of variance and/or covariance of QTL, the power and precision of the results can be greatly increased, as described in the present invention.

A special point of concern in current QTL analysis, especially in plants, is accounting for QTL-environmental (QTL-E) interactions. Few packages deal with this problem (PLABQUTL, MapQTL, and MultiCrossQTL). Instead most use a single-environment models and test the hypothesis of QTL-E interaction by ANOVA only at the final steps of the analysis (e.g., PLABQTL package). Therefore, the major advantage of multiple environment analysis of a tremendous increase in mapping resolution is lost in these tools. *Jansen RC et al., 1995, Theor. Appl. 91:33-37.* A genuine multiple environmental procedure should treat all the data collected across environments. However, such an approach is accompanied by a large number of parameters in the model (3p+1, for the simplest case, where p is number of environments), which makes it not feasible from the viewpoint of numerical optimization and statistical power. A solution to this problem proposed by *Jensen et*

al. (1995) included QTL mapping model (implemented in MapQTL) describing the effects of the target QTL and regression cofactors of co-segregation QTLs, the effects of multiple environments, and the terms of QTLxE interactions. However, such an analysis is limited to only situations where the environments can be obviously characterized by some 'physical' attributes, which is very rare in reality. The present invention solves this problem by use of a method of analysis of QTL-Environmental interaction with no limitation on the number of environments where the trait is scored.

An additional problem with already the available QTL mapping packages is the very narrow spectrum of population structures and/or designs of the mapping experiment, that are available for the user. As for the types of mapping populations, the best service is provided by a recently produced Multi-Cross QTL package which allows for such structures like BC, DH, RIL, F2, F3, full-sib families, and other regular population structures. However, it does not include such options as multiple trait analysis, selective genotyping and fine mapping. Thus, there is need for a modern user-friendly QTL mapping package which includes the described spectrum of options, as well as provides an option to complement the available population structures and experimental designs by new ones, using the shell of the present invention. No such packages are currently available. The present invention provides an option of multiple trait analysis and enables simultaneous analysis of more than two or three quantitative traits.

Numerous studies are underway in human studies to identify the target genes, the fingerprint genes and/or the pathway genes in various diseases including, but not limited to, cardiovascular diseases, the various cancers, the immune inflammatory diseases, metabolic diseases, autoimmune diseases and genetically defective

conditons. Further, the gene products of these various gene types, and antibodies to such gene products are being tested for the diagnosis and treatment of corresponding diseases. Genes may be differentially expressed in various diseases relative to their expression in normal conditions. A differently expressed gene may have its expression activated or completely inactivated in normal versus diseased conditions. Such a qualitatively regulated gene will exhibit an expression pattern within a given cell type which is detectable in control subjects, but not in the patients, and vice versa. Detectable, as used herein, refers to an RNA expression pattern which is detectable via the standard techniques of differential display reverse transcriptase – PCR and/or Northern analyses, which are well known to those of skill in the art. However, in addition to the detectable gene profile, there may be other factors such as other genes and environmental conditions that are important in the control and manifestation of a disease. For example, in breast cancer, the genes identified thus far include p53, her-2, BRCA1 or BRCA2. The present invention provides an option of multiple trait analysis and enables simultaneous analysis of several quantitative traits and functional genomics.

3. SUMMARY OF THE INVENTION

The present invention provides methods and a system designed for genetic mapping of multiple Quantitative Trait Loci, including, but not limited to, growth rate, height, weight, compositional and physiological traits, differential expression in diseases, drug resistance and environmental conditions in which the plants, animals or humans are subjected. Specifically, the traits important for plants include, but are not limited to, yield quality, productivity, resistance to pesticides, biotic stresses, suitability to modern industrial technologies, fitness related traits of wild types and

preservation of endangered species. For animals and humans, the traits include, but are not limited to, growth rate, drug resistance, disease susceptibility and simultaneous analysis of multiple traits at different stages of development and environments.

It is an object of the invention to cope with the most difficult problems of QTL mapping analysis, such as the growing number of parameters with increasing number of traits, effective QTLs and environments where the data have been collected. This resolves the basic contradiction between the real increment of available mapping information when more traits and environments are included, and the ability to utilize this information. Likewise, the method can deal with incompletely available data across the environments.

The present invention provides software for QTL mapping which fits the foregoing objectives. It is based on novel approaches and algorithms and simultaneously takes advantage of theoretical achievements of the entire world mapping community. The package or system for the multiple trait analysis described in the present invention increases dramatically the power of accuracy and precision of QTL analysis.

It is a feature of the invention to provide an interface (package) for the geneticist to employ the currently available theoretical knowledge in solving problems in a highly friendly "environment." The developed "environment" assists the experimenter in: (i) designing a new experiment; (ii) optimizing the ongoing experiment(s); and (iii) analysis and interpretation of the obtained data and results. The algorithms include genetic mapping of quantitative trait loci in plants and animals.

The system of the present invention provides a broad spectrum of analytical and modeling tools for QTL analysis, including the following functions:.

(a)   Conducting comprehensive "before-experiment" simulation analysis of a spectrum of experimental designs of QTL mapping characteristic to several types of breeding systems and organisms (outbred and inbred, with single-and several-generation data, mapping families and large populations, etc.), and helping to generate a design with optimal cost/benefit ratio.

(b)   Simulation assistance for the "ongoing experiments" conducted within a concrete project aimed to utilize the available information at any experimental step in order to optimize decision making about the next steps with respect to the individuals to be genotyped and phenotyped, markers (intervals) and traits to be further characterized. This establishes a basis for a new technology of "computer-assisted mapping experimentation," similar to classical statistical ideas of sequential experimental design, the most efficient statistical strategy. The present invention provides a useful tool for *fine mapping*, a streamlining technique in QTL mapping.

(c)   Efficient, flexible and complex (though user friendly) QTL data analysis "after the experiment," allowing for a fuller extraction of the mapping information from the data, hence higher power of QTL detection and mapping accuracy and higher benefit/cost ratio of the experiment. This "after" stage includes also large scale simulations using the "bootstrapping", empiric permutation tests and Monte-Carlo approach in order to estimate the precision of the estimated genetic parameters.

Therefore, the package of the present invention provides an "environment" in which the experimenter can: (i) design new experiments; (ii) optimize the ongoing experiment(s) by taking into account the currently available partial information; and (iii) analyze and interpret the obtained data and results. The central idea of the

package is multiple trait analysis with maximum possible flexibility in formulating and testing genetic hypotheses. The package of the present invention provides powerful simulation tools, and analytical tools. An important component of the flexibility is the possibility to include into the prepared framework new criteria of testing the significance of the detected QTLs as well as new mapping models and population structures, an option that is still not available in other packages. The proposed package of the invention provides such important options including, but not limited to, selective genotyping, accounting of 'QTL-E' interaction in the multiple environment mapping design, fitting co-segregating QTLs as regression co-factors, and allowing for linked QTLs, and fine mapping via sequential experimentation e.g. based on the "golden section" approach.

A preferred embodiment of the present invention is to provide a method to carry out the practical mapping of economically important quantitative traits in inbred and outbred plants (including trees), and in animals. Additionally, the present invention provides the option of enhanced flexibility by including additional population structures and mapping designs which make the system of the present invention an important tool for countless further theoretical development of the corresponding mapping fields. In addition to organisms of agricultural importance, mapping becomes an increasingly important tool for ecologically valuable organisms, endangered species, and species that are subject to control by human activity (like weeds, mosquitoes, many fungi, etc.). For these species, mapping QTLs of fitness components, or QTLs affecting the sensitivity to controlling agents, will involve more resources and advances in QTL analysis for multiple environments.

The method of the present invention is particularly useful for geneticists and breeders analyzing the genetic control of quantitative traits related to yield and

productivity and their numerous components, yield quality, the developmental characteristics, resistance to stressful conditions, pests, and diseases, the ability for efficient utilization of favorable conditions (e.g., fertilizers and water), suitability to modern food technology, etc. In total, these include many dozens of traits for each plant and animal species that is used for human consumption, or is a component of ecological systems dependent on human activity (like exploited forests).

The system of the present invention provides a broad spectrum of analytical and modeling tools for QTL analysis, including multiple traits analysis, QTL-environmental and QTL developmental interactions, fitting co-factors in single and multiple trait analysis, fine mapping analysis of linked QTLs, testing for dominance and epistatic effects, allows selective genotyping, significance testing with permutation and bootstrap tools and comparing alternative models, dealing with varioius types of populations (backcross, dihaploid, F2, RIL brother-sister and RIL selfing, grand daughter design, sib-mating, sib-pair analysis, etc.). The graphic user interface is user-friendly and based on Windows. The calculations are very fast. Also, tailor-made versions are possible according to needs. Another useful feature is the extensive ability to simulate data. This feature is useful for designing experiments, interpreting the results of QTL analysis and teaching the concepts of QTL mapping.

4. BRIEF DESCRIPTION OF THE DRAWINGS

The following list of the screens are enclosed by way of example to illustrate the structure of the package and the types of dialog and outputs produced during the execution of the problem solution.

Fig. 1 is a perspective view of the main menu screen of the package in accordance with the present invention, exemplifying the main menu.

Fig. 2 describe the steps of importing population data from ASCII file.

Figs. 3 describes the Population Set Define screen and its parameters.

Fig. 4 describes the Select Data Files aligning the chromosome list against the Trait List.

Fig. 5 is a view of an Input Data Report and the Commands available.

Fig. 6 describes a view of a Multi-QTL model which has detected a problem in input data.

Fig. 7 describes a view of Input Data Report in which errors are being fixed.

Fig. 8 describes a view of the Changing of the Markers Phase.

Fig. 9 describes the main menu and Project button.

Fig. 10 describes the Simulation Parameter display.

Fig. 11 describes the Chromosomes Set.

Fig. 12 describes QTL's set to display the Location in Interval.

Fig. 13 describes the Chromosome set for a number of markers.

Fig. 14 displays if there is a QTL interaction.

Fig. 15 displays the Epistas.

Fig. 16 displays the Parameters Set and the Mean values of the traits.

Fig. 17 displays the Simulation parameter for a number of genotypes-chromosomes-trait.

Fig. 18 describes the chromosome set for a number of markers.

Fig. 19 describes the QTL's set displaying a Location in interval.

Fig. 20 describes the Chromosomes Set.

Fig. 21 describes the Parameters Set.

Fig. 22 describes the Population: Backcross display for Model Parameters.

Fig. 23 describes the Population: Backcross display for Extended Parameters.

Fig. 24 describes the Population: Backcross display for Selective Genotyping.

Fig. 25 describes the Calculation Panel for d2 Models.

Fig. 26 describes the Calculation Panel for a single Chromosome set.

Fig. 27 describes the A mount model display.

Fig. 28 describes the specified Parameters.

Fig. 29 describes the Occurred Parameters.

Fig. 30 describes the Submodel for 2 traits.

Fig. 31 describes the Submodel Effect.

Figs. 33-34 describes the displays for Chromosome Information and Report.

Fig. 35 describes the status of the New Chromosome.

Fig. 36 describes the Fill Missing Markers Parameters display.

Figs. 37-39 describes the Data Check and Transformation traits.

Figs. 40-41 describes different displays for Data Check and Transformation.

Figs. 42-43 describe the Scanning Option and Scanning Parameters Setup.

Figs. 44-46 describe the Estimate option and Distribution displays for the Observed and Expected data.

Figs. 47-49 describe the significance permutation test and the Expression Editor.

Figs. 50-52 describe the significance bootstrap option and tests.

Figs. 53-56 describe the significance compare option and Comparison Tests.

Fig. 57 describes the Calculation Panel.

Figs. 58-61 describe the Multi-Simulation option and results.

Figs. 62-64 describe the print and save options.

5. DETAILED DESCRIPTION OF THE INVENTION

The proposed method (package) consists of several elements, as follows:

(1) Introduction and correction of real data. The possibility of missing data is allowed for in the system as well as some options of restoration of missing marker data.

(2) Selection of data subset to form the 'problem' (or 'job') for current analysis. One and the same data set may serve a basis for multiple specific 'problems' ('jobs').

(3) A system of "windows" that provides a high flexibility for users to choose the model of data analysis, based on the rich spectrum of algorithms built into the system:

(a) choosing concrete real data – chromosomes or specific segments, genotypes or groups of genotypes, specific traits or trait combinations, environments where the traits were scored, etc.;

(b) choosing the adequate (corresponding to the real experiment) structure of the population, the recombination mapping function, and defining the vector of genetic parameters;

(c) choosing the adequate (corresponding to the real experiment) mapping design (e.g., selective or non-selective genotyping);

(d) defining the mapping model – single or multiple trait, with a single or linked QTL per chromosome, with or without fitting cofactors from other chromosomes, for a single- or multiple-environments, etc.

The system of windows is not a closed one; it provides the flexibility for further development by including new classes of mapping problems, population structures, mapping designs, and mapping principles (for example, the interval version of the method of methods is underway and will soon be included to the package. Its major advantage is that it is distribution-free. (*Korol et al., 1984*) This is achieved by

appending corresponding parameter sets and algorithms of calculation of the employed statistical criterion.

(4) Estimation of parameters and calculation of significance. This section of the package uses the algorithms of maximum likelihood analysis to obtain the parameter estimates and to test the significance of a QTL effect in the target chromosome. Fast algorithms of optimization based on direct analytical (when possible) calculation of the derivatives are implemented. New techniques of numerical optimization for multiple parameters with single and multiple extremes will be used in the package. The system allows the user to conduct the QTL mapping analysis based on either interval or marker analysis.

This section of the package efficiently uses simulation methods to achieve its goals which include, but are not limited to, improvement the algorithms of analysis, i.e., to adapt the algorithm for the current specific situation; ability to obtain interval estimates of the parameters (confidence intervals); ability to choose the most adequate model for specific set of real data (using bootstrap permuation and Monte-Carlo analysis); and, ability to evaluate empirical threshold values for the test statistics (significance level) using permutations of the marker data relative to quantitative trait values. It is noteworthy that the last option is available for both simple interval mapping and for any interval after composite interval mapping (fitting co-factors) is conducted.

(5) Monte-Carlo analysis: This section provides broad options of simulation assistance which is important for: a) justification and interpretation of the mapping results obtained on real data; b) designing further experiments on the basis of already obtained results; c) comparative analysis of different population structures

and experimental designs; and, d) comparative analysis of different analytical approaches and different algorithms.

(6) Output of the results.   This is a system of windows for graphical and tabular representation of the obtained results.

(7) The HELP system: It is represented to enable one of skill in the art to use the system at all stages of the analysis and all sections of the entire program.

(8) Major regimes.   In general, the package provides the possibilities to conduct the QTL analysis either in a dialog regime or automatically.

(a) In the dialog regime the user can obtain the current information about the process and results of the analysis. This information can be used for optimization of the process by tuning the algorithm to the specific situation to resolve the problem. After tuning the algorithm, the user may pass to the automatic regime to conduct massive data treatment.

(b) In the automatic regime the process of analysis is conducted with a minimal participation of the user, because the analysis is proceeded according to the strategy chosen at the previous stage of dialog analysis.

(c) In both regimes the user has the possibility of interruption of the process if the user decides that the chosen regime of analysis is not successful. Then the user can pass to other options in order to try other versions of analysis of the data or to return to simulation studies, before treating the real data.

The system of the present invention is based on research achievements and simultaneous use of the major theoretical results in the field. The package of the present invention has several advantages over any of the existing software:

(i) It allows to employ multiple trait analysis, and thereby improves the biological relevance of the QTL mapping results and increase statistical power of QTL detection and mapping resolution.

(ii) Both multiple- and single-trait analysis are implemented for a broad spectrum of mapping populations, opening the possibility for applications in plant and animal genetics and breeding, with a further extension to human and medical genetics.

(iii) Because maximum likelihood (ML) analysis is chosen as a central analytical technique in the package, the possibility to take into account variance-covariance effects is allowed for by the package which gives a further increase in power of QTL detection and location precision; development of distribution-free interval version of moment-based mapping will serve an important complementation for ML analysis for cases with extreme deviation of the trait distribution from the expected normality (the graphical algorithm of scale transformation provided by the package already now solves partially this problem).

(iv) Selective genotyping design for both single- and multiple-trait analysis is for the first time implemented in the form of an interval maximum likelihood mapping QTL, taking into account the variance-covariance effect. Such a combination results in increasing the resolution power of QTL analysis.

(v) The foregoing options are implemented in the form of single-OTL per chromosome and linked QTLs as the first step of the analysis followed by multilocus mapping that is implemented as an iterative fitting of co-factors.

(vi) The method of the present invention describing the analysis of QTL-by-environment interaction, with no limit on the number of environments has heretofore,

no known analogues in the literature. The proposed algorithm can also function with massively missing data on trait scores in different environments.

The system has many exciting features. It provides a broad spectrum of analytical and modeling tools for QTL analysis, including multiple traits analysis, various types of populations (Backcross, Dihaploid, F2, RIL brother-sister and RIL selfing, grand-daughter design, sib pair analysis, etc.) analysis of linked QTLs and testing for epistasis, fitting co-factors, it allows selective genotyping, significance testing with permutation and bootstrap tools and comparing alternative models. The graphic user interface is user-friendly and has no analogues in terms of the spectrum of services and easiness of analysis. The calculations are very fast. Also, tailor-made versions are possible according to needs. Another useful feature is the extensive ability to simulate data. This feature is useful for designing experiments, interpreting the results of QTL analysis and teaching the concepts of QTL mapping. Likewise, before the real interval analysis the user has a spectrum of tools for data mining, data editing (concerning both the markers and quantitative scores), recovering missing data.

The Quick Start Tutorial (Qstart.zip), a part of the present invention, is a Microsoft PowerPoint™ presentation divided into 3 parts: Qstart1.ppt – Quickly learn MultiQTL's functions, Qstart2.ppt – Quickly load real data, Qstart3.ppt – Quickly run a simulation. The presentations demonstrates the use of the system.

The minimum system requirements for the present invention include: a Computer, Pentium Processor, Operating system Windows 95/98/2000/NT4, Hard Drive, 4 MB free, Monitor, 256-color display, RAM, 32 MB (the more the faster), and Miscellaneous Mouse and keyboard.

To install the system, the file multiqtl.zip is unzipped into a temporary directory, and then setup.exe is run. The following files are put in the directory chosen: MultiQTL.exe – the executable file, Multiqtl.hlp – MultiQTL's help file, Tip.hlp – MultiQTL's help tips, Qstart.zip – the quick start tutorial, Catalog.txt – advanced function file. Do not edit, Multiqtla.ovl – advanced function file. Do not edit, Shablon.se# - advanced function file, Do not edit, Readme.txt – Last minute updates, Users_guide.doc – The users guide (this file), Example.job – simulated example, Ex1.job – simulated example, Ex1.chr – sample data file, Ex2.chr – sample data file, Ex1.tra – sample data file, Uninst.isu – uninstall information, Three ActiveX files are placed in the Windows/System directory, Msflxgrd.oca   Msflxgrd.dep   and Msflxgrd.ocx.

The system may be started using one of these three options: 1. Press the Start button, choose Programs and click on the MultiQTL icon; 2. Click the MultiQTL icon on the desktop, and 3. Double click on the multiqtl.exe file that is in the directory you installed MultiQTL. The default directory is:

C:\Program Files\MultiQTL\MultiQTL\

The Multiqtla.ovl and catalog.txt must be in that same directory for the program to run. The shablon.se# file is needed for printing to a *.pcx file.

The system may be closed in the same way as most Windows programs. To Exit The system, one of the following tasks is performed: 1. Select Exit from the main menu, 2. Click the main window's Close button (the 'X' on the upper left corner) and 3. Press <Alt> + <F4>.

FIG. 1 describes MultiQTL's main window. At the start of the program, MultiQTL's logo shows up for a second, disappears automatically and the main window appears:

The main window is the workspace, and it contains the tools to load, create, edit and print QTL data. The Main window consists of (from top to bottom):

Title Bar -The Title bar shows MultiQTL's icon, a title relevant to the active window, and the minimize, maximize and close buttons (not shown).

Menu Bar- The Menu bar displays the menu headings. The menus available depend on what type of window is active in the workspace.

Toolbar- The Toolbar provides fast access buttons to frequently accessed commands. If a command is not available, the button appears greyed-out. To add or remove the Toolbar, use the View menu.

Workspace- The workspace is the area of the main window where the user work and see all the active windows.

Status Bar- The Status bar, which appears at the bottom of the main window, contains a brief description of the item the user is pointing at (not shown).

The system (also referred to as "MultiQTL") may be taken off from a computer using the following procedure:

1. Click the start menu button to open the start menu; 2. Point to settings, and then click Control Panel. The Control Panel opens; 3. Double-click the Add/Remove Program icon. The Add/Remove Programs Properties dialog box opens; If necessary, click the Install/Uninstall tab to bring it to the front; 4. Scroll through the list box to find 'MultiQTL'; 5. Highlight 'MultiQTL'; 6. Click the Add/Remove button. The Uninstall setup program starts; and 7. Follow the on-screen prompts.

For more background on the scientific foundation of MultiQTL, press the yellow question mark in the Toolbar or About MultiQTL from the Help menu. Press the Publications button to see a list of papers by Korol and co- authors. See also Appendix A.

6. PREFERRED EMBODIMENTS

EXAMPLE 1. LOADING REAL DATA

In order to analyze QTL data, the data files have to be prepared in special formats. All the files should be in ASCII format, and arranged as described herein.

Chromosome File Formats (*.chr and *.mrk)

The chromosome data files must have a '*.chr' extension. The data is arranged in a matrix with each row having a unique marker name and then a row of numbers or symbols representing genotypes. The default settings are that 0 represents a missing genotype, 1 represents an 'AA' genotype, 2 represents an 'aa' genotype and 3 represents an 'Aa' genotype. In the F2 population, 4 represents a dominant 'A' maternal genotype and 5 represents a dominant 'a' paternal genotype. The default settings can be overridden by any other symbols while selecting the data files through the 'Select Data Files' window (see page 12). All the rows should have the same number of genotypes. The file appears as follows:

```
mar1a 1 3 2 3 1 1 2 1 3 2 1 1
mar2a 1 0 2 3 1 1 2 1 0 2 1 1
mar3a 1 2 2 3 1 1 2 1 1 2 1 1
mar4a 1 2 3 3 3 1 2 3 1 2 3 1
mar4b 0 0 0 0 3 1 0 0 1 0 3 1
m6a 3 3 0 0 3 1 1 0 1 3 3 1
m6b 3 3 1 2 3 1 2 2 1 3 3 1
```

m6d 3 3 1 2 3 1 0 0 1 3 3 1

There doesn't have to be a space between the numbers (i.e. Marker_name 02 103 2 is OK), but there must be a new line character ('Enter') at the end of each line. The '*.mrk' data files are exactly the same as '*.chr' files, except that they have only one marker data in them and are therefore only one line long. There should be exactly the same number of genotypes in that line as the number of genotypes in the *.chr files. The file should appear as:

ma4 2 2 3 1 3 0 1 3 1 3 0 1 1

Each file corresponds to one marker, and it is used to add new markers to a calculation. An example chromosome file – 'ex1.chr', is included in the MultiQTL directory.

Any error in the data will be shown in the 'Error Report'. Special editing tools are available to fix the data.

Trait File Formats (*.trt and *.tra)

The trait data files must have a '*.trt' extension. The data is arranged in a row of trait values. The number of values are exactly the same as the number of genotypes in the *.chr files. By default, the symbol $ represents a missing trait value. The default settings can be overridden by any other symbols while selecting the data files through the 'Select Data Files' window.

The file should look like this:

54.5 68.5 87.5 88.5 $ 88.5 82.5 78.5 88.5 80.5 95.5 75.5

Any error in the data is shown on the 'Error Report'. Special editing tools are available to fix the data.

To simplify the process, a '*.tra' file can be made that contains all the traits values. The data is arranged in a matrix with each row having a unique trait name and then a row of trait values. All the rows should have the same number of values (exactly the same as the number of genotypes in the *.chr files). As in *.trt, by default the symbol $ represents a missing trait value. The default settings can be overridden by any other symbols while selecting the data files through the 'Select Data Files' window. The file should look like this:

Bur 54.5 68.5 87.5 88.5 $ 88.5 82.5 78.5 88.5 80.5 95.5 75.5

Yg2 4.28 4.10 6.45 2.25 11.25 12.55 $ 8.90 10.14 8.83 13.45 9.14

Multiqtl.exe takes a *.tra file and creates from it multiple *.trt files. Each row (trait name and values) becomes a separate trait file in the *.trt format. Each file gets a unique name that comes from the name of the trait (Bur.trt, Yg2.trt, etc.). The program also checks that there is exactly the right number of genotypes in each row.

Families (or Environmental Groups) File Format

(*.env)

If the user chooses a population with the 'groups' (families) option, then there should be a *.env file that divides the genotypes into groups. If there are 117 genotypes that are divided into 4 groups, the first group holding 32 genotypes, the second group holding 31 genotypes, the third 26 and the fourth 28, then the file should like this (omitting 24 characters from each row):

0.......0   0   0   0   0   0   0

1.......1   1   1   1   1   1

2.......2

3.......3   3   3

Information File Format (*.inf)

A user creates an input information file when importing data from an ASCII file. The file, *.inf, contains information on the number of genotypes, population type, usage of selective genotyping and families, and a list of chromosome and trait files.

Automatic Data Checking

A user automatically checks the data for errors. If errors are found, tools for fixing the data are available and an error.txt file is written with a list of all the errors.

If there is an error in the *.chr or *.trt files (wrong data symbol or wrong number of input data units in a row) a window opens with a summary of all the errors, and editing tools for fixing the data are provided.

If there are intervals with a recombination rate of more than 50% due to an erroneous designation of marker alleles or because of repulsion phase of markers, an automatic fixer will open and allow the user to fix the data. New chromosomes are written, and they are used for the data analysis.

Loading Real Data

Select 'Data/Import/from ASCII file' from the main menu to load the prepared data files. FIGs 2 & 3 describe the 'Main Menu' and the 'Population Set Define' window, respectively. The procedure is described below.

Choose a *.tra file and press <Open>. MultiQTL will take the *.tra file and split it into many one lined *.trt files.

Select a population and enter a name of a population set into the 'Population Set Define' window. An *.inf file is created (in this example, the file is named qaz.inf). Click <OK>.

In the 'Select Data Files' window, select the chromosome files from the 'Chromosome List' and the traits from the 'Trait List'. To select multiple files, use the <shift> and <control> keys with the mouse buttons, or drag the mouse over the list while pressing the left mouse button. Next, change the coding symbols to match the codes of the input files. Click <OK>.

FIGs 4 & 5 describe the 'Set Data Files'.

If there are no problems, the procedure is continued as described below at the model creating stage. If there is an error in the *.tra file (the number of input data units in each row is not identical), an error window will notify it, and a fixing tool is provided. FIGs 6, 7 & 8 describe the 'Input Data Report'.

Fix the file by selecting a command (edit cell, insert cell, delete cell or split row) and then selecting a cell to perform the command on it. Click <Save> and <Finish> to continue. Now start loading the data again.

If there is an error in the *.chr file (there is an unidentified symbol or the number of input data units in a row does not equal the number of trait values), a summary of the errors appears and a fixing tool is provided.

Click <Next> to edit the data. The chromosome data editor works the same as the trait data editor.

Click <Save> and <Finish> to continue. Now start loading the data again.

If in a chromosome there are intervals with a recombination rate of more than 50% due to an erroneous designation of marker alleles or because of repulsion phase of markers, MultiQTL will notify about the problem and will provide a tool for fixing it. The fixed chromosome is named the same as the original chromosome, trailed by a '%1' (i.e. ex1.chr   ex1%1.chr, ex1%2.chr, etc.). The original

chromosome does not change. The new chromosomes is saved, and is used for the data analysis.

FIG 9 describes the 'Changing of the Marker Phase' window.

Select the markers that have their phase changed by clicking on of the radio buttons, and press <Apply>. Click on <Save and Exit> to continue. MultiQTL will notify the user about the newly created chromosomes.

The user has now reached the model creating stage.

EXAMPLE 2

<u>CREATING SIMULATED DATA</u>

In addition to real data analysis, there is also an option to create data that simulates real data. This option is useful for comparing different mapping designs, interpreting the results obtained with real data and teaching the principles of QTL mapping and analysis. To start the simulation, choose Simulation from the Project menu or click the Simulation button on the toolbar.

Figure 10 describes the 'Multi QTL' window.

To simulate data, the following parameters of the simulation are entered:

1. Number of genotypes (no limitation); 2. Number of chromosomes (maximum of 30); 3Number of traits (maximum of 3 in this version); 4. Type of mapping population (Cross type); 5. Mapping function; and 6. Percent of lost markers (to simulate missing data).

Figure 11 describes the 'Simulation Parameter' window.

When there are missing markers, the user can either leave them missing or restore them. When missing markers are not restored, the results of the analysis are less significant. If they are restored the results may be more significant, but they are

less accurate. The 'Lost markers percent' entry limits the percent of markers that can be missing. Any more missing markers are automatically restored.

After entering all the parameters, press <OK>.

Simulating Backcross, RIL, Dihaploid, etc.

First the user can simulate a Backcross population with 3 chromosomes and 2 traits. The simulation procedure is the same for all the other populations except for F2.

Setting Marker Locations - Figure 12 describes the 'Chromosome Set' window.

To locate markers on the chromosome, the following procedure is followed: 1. In the 'Chromosome Set' window, select a chromosome by marking the circle on its right (or <All> the chromosomes together); 2. Enter the 'Number of markers' and the 'Chromosome length (cM)', for that chromosome. The length is in centiMorgan units. The markers at both ends are counted; 3. Press <Set markers>. The markers are shown on the chromosome in equal distances; 4. Move any marker by dragging it with the left mouse button. The chromosome has to be chosen to move a marker on it; 5. Delete any marker by clicking on it with the right mouse button. The chromosome has to be chosen to delete a marker on it; 6. Repeat 1-4 for all the other chromosomes; 7. Continue editing the markers until satisfied; and 8. Press <Next>.

Setting QTL Location -Figures 13 and 14 describe the 'QTL's set' window and 'Chromosome set' window respectively.

To map QTLs on the chromosome, the following procedure is followed: 1. In the 'Chromosome Set' window, click on a chromosome; 2. Fill in the parameters (Location and Effects) in the 'QTL's set' window; 3. Press <OK>; 4. Repeat for as

many QTLs as you like. There can be many QTLs on a chromosome but only one QTL per interval; 5. To delete a QTL, right-click on it; and 6.Press <Next>.

Setting Epistatic Interactions - Figures 16 and 17 describe the windows used for setting epistatic interactions.

To add the epistatic effect of two interacting QTLs, the following procedure is noted: 1. Answer <Yes> to the 'QTL interaction' query; 2. Click on any two QTLs; 3. Enter their 'Epistasis'; 4. Press <Add Pair>; 5. Repeat for any two QTLs. In the current version, each QTL can interact with only one other QTL; 6. Delete any pair you want; and 7.Press <Next>.

Trait Distribution Parameters – Figure 18 describes the 'Parameters Set' window. To enter the distribution parameters for the traits, the following procedure is followed by entering: 1. Mean values; 2. Heritability; 3. Residual correlation; and 4. Press <OK>.

EXAMPLE 3

Simulating F2

The following procedure describes how to simulate a F2 population with four chromosomes and three traits. Figure 19 describes the 'Simulation Parameters' Window. Press the simulation button or press Simulation from the Project menu, and enter the appropriate data.

Press <OK>.

Figure 20 describes the 'Chromosome set' window.

The markers may be placed on the chromosome using the following procedure: 1. In the 'Chromosome Set' window, select a chromosome by marking the circle on its right (or <All> the chromosomes together); 2. Enter the 'Number of markers' and the 'Chromosome length (cM)', for that chromosome. The length is in

centiMorgan units. The markers at both ends are counted; 3. Press <Set markers>. The markers are shown on the chromosome in equal distances; 4. If needed, move any marker by dragging it with the left mouse button. The chromosome has to be chosen to move a marker on it; 5. Delete any marker by clicking on it with the right mouse button. The chromosome has to be chosen to delete a marker on it; 6. Repeat 1-4 for all other chromosomes; 7. Continue editing the markers until satisfied; and 8. Press <Next>.

Setting Dominant Markers – To set dominant markers, the following procedure is used: 1. Click on a marker. A red-blue box will appear on the marker; 2. Carefully click on the red box to mark the marker as dominant-maternal, or on the blue box to mark the marker as dominant-paternal; 3. Repeat 1-2 for all other dominant markers; and 4. Press <Next>.

Setting QTL Location (F2) – Figures 21 and 22 describe the procedure to map QTLs on the chromosome as: 1. Click on a chromosome to place a QTL; 2. Fill in the parameters (Location and Effects) in the 'QTL's set' window; 3. Press <OK>; 4. Repeat for as many QTLs as you like. In the current version there can be only one QTL per interval; and 5. Press <Next>.

Trait Distribution Parameters (F2) – Figure 23 describes the 'Parameter set' window to enter the parameters for the traits as: 1. Mean values; 2. Heritability; 3. Residual correlation; and 4. Press <OK>.

EXAMPLE 3

MODELS

To analyze the data and compare different hypotheses a user creates modelsand submodels, using the data available. Models differ in the number of traits

or the number of QTLs in them and submodels differ from the main model by nullifying (lowering to zero) some of the effects in the model.

A vector of parameters defines a hypothesis. This vector is regarded as a model. Different parameters define different models. The differences from one model to another can result from: 1. Difference in traits; 2. Difference in number of QTLs (e.g. a single trait vs. two traits); 3. Different mapping functions; 4. Difference in selective genotyping (difference in trait or tail size).

Each change can be saved as different model. All the models can be analyzed and compared to see the effect of the changes on the LOD;

Figures 24, 25 and 26 describe the 'Population: Backcross' window.

The following procedure is used to create a model after loading or simulating data: 1. If the Model building window is closed, select Create from the Model menu; 2. Enter a model name (has to be unique); 3. Select the number of traits; 4. Select traits from the 'Names of Traits' list; 5. Select a mapping function; 6. Select the number of QTLs (single QTL or two linked QTLs); 7. The 'Initial Submodel Default' is by default already checked on. The effects of changing this option are covered on page 25; 8. The 'Selective Genotyping' check is off by default. Checking this option requires selecting a trait ('Selected Trait') and adds a tab called 'Selective Genotyping', for specifying the parameters of selective genotyping; 9.Click on the 'Extended Parameters' tab; 10. Select the 'Calculation Method': 'Marker Analysis' or 'Interval analysis'; 11. Select 'Marker Restoration' if you want MultiQTL to restore the missing markers, or select 'Ignore Marker Loss' to ignore the missing markers; 12.'Selective genotyping' can be selected only when the data was simulated (in real data mode the option is disabled). Click on the 'Selective genotyping' tab, and slide the right and left tail cutting bars, to define the tail individuals to be genotyped for

marker loci. There must be at least 50 objects left on each side to continue. The sliding bars sensitivity is controlled by choosing 'Rough Tuning' or 'Fine Tuning'; 13. To add the model Return to the 'Model Parameters' window by clicking on its tab, and then click on <Add Model>; 14. To add more Models, change the settings of the model and then click on <Add Model>; 15. Before clicking <OK>, scroll through all the models and change or delete them; and16.When finished, press <OK>.

Fitting Model Parameters

A user can define 4 models:

Model name	Number of Traits	Number of QTL
d1	1	1
d2	2	1
d3	1	2
d4	2	2

Figures 27 and 28 describe 'Calculation Panel' windows. The procedure is: 1. If the Calculation Panel window is closed, select Open/Single from the Model menu. The Calculation Panel opens; 2. Select a model from the 'Models' option, by clicking on the arrows; 3. Select the graph buttons you want to calculate by one of these methods: (a) Press <All> to calculate all the graphs, (b) Press on a trait button to calculate all the graphs of that trait, (c) Press on a chromosome button to calculate all the graphs of that chromosome, (d) Press one graph button; 4. To deselect a button, press it again; and 5. Press <Compute>.

When a graph is calculated, it is shown as a thumbnail on the button itself, with the maximum LOD of the graph written on the top. Press <About Model> to see more information on the specific model.

The <About Simulation> and <About Families> give more information, but only for simulated data or models with families, respectively.

Figures 29, 30 and 31 describe the windows for 'About model' 'specified', and 'Incurred'.

To Enlarge a graph simply click on it. Notice that the LOD maximum on the thumbnail turns red to show that this graph has been enlarged since last calculation.

If a QTL was simulated on an interval, a small green triangle will be placed on that interval. Clicking on it gives more information on the QTL.

It is noteworthy that the d1 model is a model with single trait analysis, but the d2 model is of two-trait analysis. In a different model of MultiQTL there is real multiple trait analysis also available.

Models d3 and d4 are three dimensional models of two-linked QTLs. Model d3 is a single trait analysis model, but d4 is a two-trait analysis model. It takes more time to calculate these models because of the sophisticated 3D calculations.

Submodels

Submodels are the same models of the same data but with different calculating assumptions. The effect of a trait or the variance and covariance effects can be forced to zero. This tool helps in calculating the significance of an effect.

In the simple case of a one-trait model with the Backcross population, there can be 4 different submodels; for example: 1. Both the main QTL distribution effect and the variance effect (allowing for different residual variances in the QTL groups) are calculated; 2. No variance effect. Only the main effect is calculated; 3. Effect=0. Only the variance effect is calculated; and 4. Both the effect and the variance effect are not calculated. This option is trivial because the LOD will be 0.

In more complicated situations, like a two-trait analysis of two-linked QTLs, a user can have many more submodels.

Figures 32 and 33 describe a 'Submodel'.

To create a submodel, the following procedure is used: 1. Open a LOD graph by clicking on a graph thumbnail; 2. Select Add from the Submodel menu; 3. Choose the effects that you don't want them to be calculated; 4. Press the <OK> button. The LOD graph of the new submodel will appear in a new color. In single trait analysis all the submodel LOD graphs will appear on the same LOD graph window, and in multiple trait analysis the graphs will appear in different windows and 5.Repeat the process for all the submodels needed.

Each time a user presses <OK> on the Submodel window a submodel will be created, unless an exact copy was already declared.

EXAMPLE 4

CHROMOSOME EDITING

The system offers tools for editing the chromosome and trait input data. The chromosome-input data can be edited only if it is real data. The trait-input data can be edited for real and simulated data.

Figures 34-38 describe the windows for 'MultiQTL', 'Chromosome Information', 'Changed Chromosome Report', 'Status of the new chromosome' and 'Fill missing markers parameters', respectively.

To edit the chromosome data, the following procedure is used: 1. Choose 'Chromosome edit' from the Data menu or Click on the 'chromosome edit' button in the toolbar menu; 2. Select a chromosome to edit from the chromosomes list; The following information is given on the chromosome: Marker name and number, defined objects on marker, interval length and defined objects on the interval; 3. Select an editing operation to perform by clicking on a radio button in the 'Change markers' options; 4. To delete a marker, click on the 'Select marker to delete' radio button. Now click on a marker to delete it. The new chromosome without the deleted

marker will be shown in a new window. After reviewing the changes, close the 'Changed Chromosome Report'. To finish the operation, select the status of the new chromosome. Click <OK> when finished; 5. To fill missing markers, click on the 'Fill markers' radio button. Enter the requested parameters and click <OK>. A window with the new chromosome will appear, and the user is prompted to select the status of the new chromosome; 6. At any time the user can look at all the changes made to the chromosome by clicking on the <History> button; and 7. While looking at the history window, the user can undo all the changes by clicking on the <Undo> button. Click on <Apply> to apply the changes.

EXAMPLE 5

<u>TRAIT TRANFORMATION</u> – The trait transformation tool is available for real and simulated trait data.

Figures 39-43 describe windows for 'Multi QTL', Data Check and Transformation, 'Left Individual', 'Data Check amd Transformation', 'Left Individual', 'Data Check and Transformation,' and 'Data Check and Transformation', respectively.

The procedure to view the distribution of the traits values and transform them is: 1. Choose Trait Transformation from the Data menu or click the Toolbar icon; 2. Select the number of traits from the 'Traits' list; 3. Select the traits to transform from the 'Names of traits' list; 4. Select the analysis method: 'By marker' or 'By interval'; 5. Select a chromosome from the 'Chromosome name' option; 6. Change the 'Marker/Interval No.' by clicking on the arrows; 7. Change the number of groups in the distribution function by clicking on the arrows (above the <Print> button). The default is 30; 8. Transform the graphs by dragging the 'Scale Transformation' bar right or left; 9. Cut off tails from the analysis by dragging the 'Tails Cutting' bar to the

right; 10. Click on a red tail to see the properties of the points in the tail, 11. Check the 'Spline' box to smooth the graphs; 12. Replace the existing data with the transformed data, or add the transformed data as an additional trait. The <add> button is operational only when using real data; and13. Reset the data to start transforming again or press <Cancel> to exit this window.

On the right side of the window we can compare the values of the 'Observed' data and the 'Converted' (transformed) data for these parameters:

1. M – average value of the trait.

2. S2 – standard deviation of the trait value.

3. Cv – coefficient of variation (in percent).

4. A – coefficient of asymmetry.

5. E – coefficient of curtosis.

To view the joint distribution of two traits by marker-defined (or interval-defined) groups as a scatter diagram: 1. Click the 'Two traits' radio button; 2. Select a chromosome from the 'Chromosome name' option; and Change the 'Marker/Interval No.' by clicking on the arrows.

The 'Replace' and 'Add' options are not operational in this version.

The 'Scale transformation' and 'Tail Cutting' sliders also don't work yet.

To see the data without the variance, check the 'No var.' Checkbox.

EXAMPLE 6

ANALYZING THE RESULTS

Scanning - The first stage of the data analysis is to fine scan the LOD graph. Figures 44 and 45 describe the scanning windows. The procedure is: 1. In the 'Calculation Panel', open a LOD graph window by clicking on a graph thumbnail (LOD graph taken from example1.job: two-trait single QTL simulation); 2. Select

Scanning from the Scanning menu; 3. Enter the scanning parameters for the scan. The default parameters usually give a satisfactory result; and 4. Press <OK>.

1. The Scan option DOES NOT change the data. It only shows it in finer detail on the screen.

2. To Unscan the graph (return to the original coarse view), select Unscanning from the Scanning menu.

3. Submodel graphs can be added (Submodel/Add menu), and each graph can be scanned. To scan a graph when multiple submodel graph are shown, select a graph by clicking the radio button corresponding to the appropriate graph, deselect other selected graphs, and then go to stage 3 and scan the graph.

4. In two-linked models, the scanned three-dimensional graph is shown.

Estimate - The estimation table shows a table with the numeric results of fitting the model, from one analysis run of the data. This table gives approximate estimates of the values of the data fitting analysis.   Figure 46 describes the 'MultiQTL – Model d2 Chromosome' window.

To see the estimation table: 1. Click on a thumbnail graph in the calculation panel; 2. Select a graph (if there is more then one graph); and 3. Click on the Estimate menu.

The model is calculated and fitted once and the estimation table appears with the following data (left to right): 1. Interval number (Interval) in interval analysis and Marker number (Marker) in marker analysis; 2. Interval size (Size) in interval analysis. Not in marker analysis. The size is in centiMorgan units; 3. Interval length in centiMorgan units. Not in marker analysis; 4. Highest LOD in interval (LOD); 5. Number of objects genotyped for the flanking markers (nObject); 6. Average value of the trait (aver.(m)); 7. Size of effect (eff.(d)); 8. Residual standard deviation; 9.

Columns 6,7,8 are repeated for all the traits that are analyzed; 10. Correlation

coefficient for the traits (when more than one trait is analyzed); and 11. In the F2

population models there may be another field titled 'h' for dominant effect.

The table shows the data for each chromosome. Highlighted in blue are the

intervals with local maximum of LOD score, and highlighted in red are the global

markers.

In two-linked QTL models the estimation table has the same data for every

pair of intervals from the two traits.

Distribution

To see how the putative QTL explains the observed distribution of the trait:

1. The user chooses a submodel and

2. Select Distribution from the main menu.

The user gets a distribution window with the trait distribution in the QTL

groups. The user can change the number of histogram intervals in the graph, and

you can go through all the chromosome intervals, to see the distribution in each

interval.

Figures 47 and 48 describe the 'Distribution' windows.

The green and yellow lines are distribution graphs for each QTL group

separately, and the blue line is their sum. This blue line is what is expected, but the

observed distribution is the red line.

Press <Cancel> to exit the Distribution window.

In a two-linked QTL model, the user can see also the distribution of the

epistasis effect.

Permutation Test

The permutation test is a method for testing significance thresholds of QTL LODs. To check the significance of the putative QTL effect(s), run a permutation test. The test runs the data with many permutations, and shows how many times we get a higher LOD through a permutation, compared to the LOD obtained with the initial data. Figures 49, 50 and 51 describe the different Permutation Test windows. To run the permutation test, the following procedure is used: 1. Choose a model or submodel; 2. Select Permutation from the Significance menu; 3. Enter the number of runs tests (default is 1000); 4. Enter the Critical LOD Value. This is the value the test tries to overcome. The default is the maximum LOD of the chosen model; 5. Press <Start>; and 6. T o stop the test press <Close> or <Stop>.

After the test stops, an advanced permutation test can also be performed: 1. Click <Advanced>; 2. Click <New Expr>; 3. Enter a new expression using the 'Expression Editor'; and 4. Press <Apply>.

The significance score is shown above the significance bar.

The <reset> button resets the permutation count. Each test is done 1000 times (or any other 'Perm. Number' defined by the user).

The filter option enables the system to ignore intervals that are estimated to have low LODs, and thus increase calculation speed. When the filter equals 20%, all the intervals that their estimated LOD is lower than 80% of the maximum LOD are not calculated. We advise beginner users to keep the filter at 100% that means that the filter is unused.

The 'Advanced Permutation Test' takes into account only the permutations that were calculated in the basic permutation series, and not any that were ignored by the filter.

Bootstrap Analysis -

The Bootstrap test takes the analyzed data, and constructs from it a random set of data by randomly resampling data elements from it. Each time a data set is randomly chosen we are virtually performing an experiment. This can be done many times, and it gives us valuable information on the size of the experiment that we have to perform to get significant results. Figures 52-54 describe the Bootstrap Test windows.

To perform the Bootstrap test, the following procedure is used: 1. Choose a model or submodel; 2. Select Bootstrap from the Significance menu; 3. Enter the number of samples (default is 1000); 4. Press <Start>; and 5. To stop the test press <Close> or <Stop>.

The results are shown on a 2D bar graph with the X-axis as the intervals and the Y-axis as the proportion of runs with the highest LOD to be in that interval.

The results are summarized in a table that shows the mean value and standard deviation of the intervals with the maximum LOD in all the bootstrap tests.

The filter works as in the permutation test, and is not recommended in the beginning.

In two-linked QTL models the bootstrap test results are shown on a three-dimensional graph. The intensity of the colors shows the amount of times the maximum LOD was calculated in a specific cell. Clicking on a cell, after the test stops, will show the exact results. The table is the same as before, but with columns for the added effects.

Submodel Comparison -

To check the significance of a specific effect, the user can compare two submodels that differ with respect to this effect. The comparison algorithm takes the results of the submodel with the smaller number of parameters (hypothesis H0) and

simulates data many times. For each simulated set, the second model (hypothesis H1) is fitted. The user checks how frequently the maximum LOD values corresponding to H1 but using simulated data based on H0 will be higher than the real LOD of the submodel corresponding to H1. Only two submodels where one is a submodel of the second can be compared in this way. Two submodels that don't fully overlap cannot be compared. Figures 55 and 56 describe the comparison test windows.

To compare two submodels the following procedure is used: 1. Select two submodels; 2. Press Compare/Submodels from the Significance menu; 3. The Comparison window opens; and 4. Press <Start>.

In the example, the variance effect is not significant. This effect is the sole difference between the two compared submodels, and the comparison showed no significant difference).

Models Comparison

To check the significance of a specific model, a user can compare two models (e.g. single QTL vs. two linked QTLs). The comparison algorithm takes the model with the smaller number of variables and simulates it. It then checks to see if the maximum LOD of the simulated model is bigger or smaller then the LOD of the second model. Only two models of the same traits on the same chromosome can be compared. Figures 57 and 58 describe the 'Comparison Test' windows.

To compare two models the following procedure is used: 1. Open two graph windows with different models. Select two models, one from each window. If there is only one model in each window, it is automatically chosen; 2. Press Compare/Models from the Significance menu; 3. A message box appears asking you to select one option: <Yes> will compare the two chosen models, <No> will deselect

the two-linked QTL model and <Cancel> will deselect the single QTL model. Select <Yes> after selecting the appropriate models; 4. The Comparison window opens; and 5. Press <Start>.

In the example, a user can see that the two-linked QTL model is not much more significant than the single QTL model. In 13.57% of the tests, the LOD of the single QTL model was larger than the LOD of the two-linked QTL model.

EXAMPLE 7 MULTISIMULATION

Multisimulation

The Multisimulation option allows checking significance of existing effects, calculating the explained variability of these effects, and the adding of specific effects to check if they are significant and what variability they can explain.

Figures 59-63 describe the 'Multisimulation' windows. If an effect is already present in a model then the stages of creating a Multisimulation and checking for significance are: 1. Open a Calculation panel with the models thumbnail graphs; 2. Check the radio button of the 'Multiple' option in the 'Chromosome Set' box. The Multisimulation panel is added; 3. Select the thumbnail graphs for the Multisimulation by clicking on them. All the selected graphs must be of the same trait; A second on a graph deselects it; 4. Click on <Add set> and enter a unique name for it. Click <OK>; 5. Click on <Open> and the graphs of the selected thumbnail graph will shown in a single panel. Scroll bars will be available when needed; 6. Click on MultiSim from the main menu. The 'Explained variability' window opens; and 7. Click on <Start> and the simulation of all the models is calculated simultaneously; 8. When the simulation stops the percent of explained variability and distribution, with the Heritability will be shown. You can switch between the traits and between levels of significance; 9. Clicking on a radio button below one of the graphs will open a window with a table of

the results of the Multisimulation of that model; 10. Clicking on a QTL Symbol (green triangle) in a graph will show the QTL properties; and 11. In Two-linked QTL models, the same options are available. Click on any cell to see the exact percentage of maximum LOD hits in that cell. Checking a radio button will open a window with the results in a table.

EXAMPLE 8

FINE MAPPING

Special analytical tools, statistical genetic models and complex experimentation are currently required to conduct QTL mapping . The present system and method can be modified by one skilled in the art to carry out fine mapping techniques which take into account the variance and covariance effect of the analyzed QTL.

EXAMPLE 9

CO-FACTOR ANALYSIS

The present invention provides a Fitting Co-Factors algorithm which reduces the background noise and the accidental influence of QTLs from the chromosomes. The procedure involves sequential chromosome analysis of QTL presence while subtracting the effects of the QTLs from other chromosomes. Fitting –cofactors may therefore discover QTLs that cannot be detected the usual way. The algorithm used is iterative since the placement and the effects of the QTLs are unknown in advance. First, the most powerful QTL is found and its influence on the other chromosomes is nullified. Second, the next most powerful QTL is searched on the remaining chromosomes and its effect is nullified, and so on. This procedure is repeated until no QTL are found on the remaining chromosomes. Third, the fitting of the QTLs is performed by specifying the QTLs (without any influence from other QTLs) in the

order of their power. This procedure is repeated over and over again until the difference between the parameters of each QTL on two consecutive iterations is less than a pre-selected value.

EXAMPLE 10

SERVICE FUNCTIONS

PRINT

While working with the system, a user has the option to print all the windows and graphs that the user sees. Either the user will have a <Print> button in the window, or the user can click Print in the Service Options menu (<Ctrl>+<p>also works).

The user can select between printing to a printer or to a file in the PCX format that is readable by almost all the graphic programs.

The file will be saved in the user's data directory. The user must have the file shablon.se# in that directory. The user can also change some of the fonts by clicking the Change Font radio button.

SAVE

The user can save all the calculations with the Save Job option in the Service Options menu. The saved file will have a job extension. To save in a different name select the Save Job As... option.

NOTEBOOK

Throughout working with the system, the user can open a virtual notebook to write in it whatever the user wants. Just click Notebook from the main menu and write in the notebook. When finished, click Save & Exit.

It will be appreciated by persons knowledgeable in the art that though reference was made to plants and animals, it is not limited thereto. For example, the application is also applicable to the mapping of human genes.

The scope of the described invention is intended to include all embodiments coming within the meaning of the following claims. The foregoing examples illustrate useful forms of the invention, but are not to be considered as limiting its scope, as those skilled in the art will readily be aware that additional variants and modifications of the invention can be formulated without departing from the meaning of the following claims.

It is also to be understood that the following claims are intended to cover all of the generic and specific features of the invention herein described, and all statements of the scope of the invention which, as a matter of language, must be said to fall there between.

CLAIMS

What is claimed is:

1.   A method for mapping quntitative trait Loci in a plant or animal

population, said method comprising the steps of:

(a)   selecting one or more traits across said population, thereby

providing at I.A method for mapping quantitative trait <u>Loci</u> in a plant or animal

population, the method comprising: east one quantified population phenotype;

(b)   identifying at least one genetic marker associated with the

distribution of the trait, and

(c)   selecting the genotypes and markers defining one or more

specific traits according to predetermined parameters.

2.   The method of claim 1, wherein the plant or animal population includes:

(a) species with variuos reproductive systems including selfing and

outbreeding;

(b) species with various types of life histories including annuals, perennials,

diploids, polyploids or trees; or

(c) species of agricultural, industrial and ecological importance.

3.   The method of claim 1, wherein the traits include:

yield, components and , quality;

resistance to abiotic and biotic stresses;

anatomical, physiological and biochemical characteristics;

quantitative expression patterns of different genes; and

traits with quantitative threshold manifestations.

4.   The method of claim 1, wherein the mapping of the quantitative trait

Loci is carried out in one or more environments.

5.   The method of claim 1, wherein the mapping of the quantitative trait Loci is carried out during one or more stages of development of the population.

6.   The method of claim 1, wherein the mapping of the quantitative trait Loci is carried out in one or more families of the population.

7.   The method of claim 1, wherein the mapping population includesdifferent population structures including backcross, F2 or F3.

8.   The method of claim 1, wherein the mapping population includes phenotypic data, or marker data from one or more generations.

9.   The method of claim 1, wherein the mapping method includes: the multilocus analysis of correlated trait complexes; the allowance for variance and covariance effects; the discriminating between linkage and pleiotropy; the testing for dominance and over dominance; and the detection of epistasis.

10.   The method of claim 1, wherein the method includes: the marker and interval analysis, the maximum likelihood method, the method of moments or the method of regression analysis.

11.   The method of claim 1, wherein the association of the trait and the genetic markers is determined by applying a statistical model or tools for testing of significance and evaluation of precision.

12.   The method of claim 11, wherein the tools for testing of significance and evaluation of precision include: the permutation tests for the lod values and user defined criteria; the bootstrap analysis ; the Monte-Carlo simulations; or comparisions between different method and versions of the same method.

13. The method of claim 1, wherein the pre-determined parameters include at least one of a group consisting of:

a single or multiple trait,

a single or linked QTL per chromosome, with or without fitting cofactors from other chromosomes, and
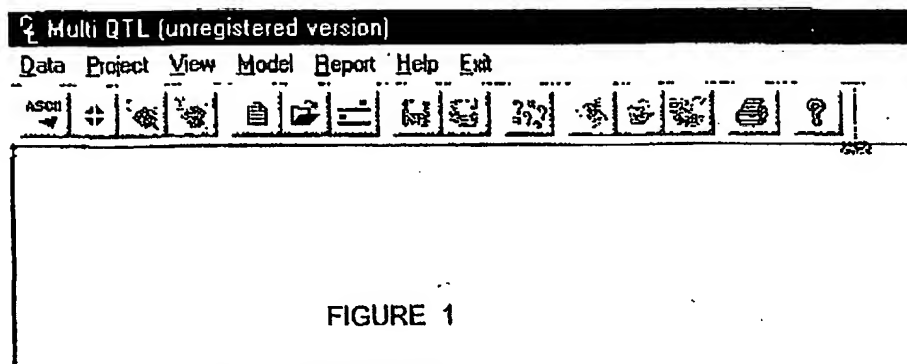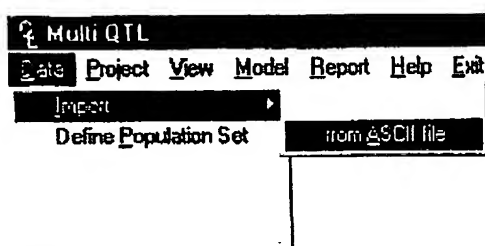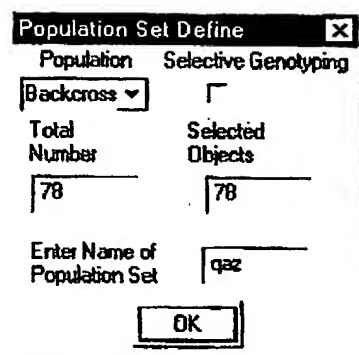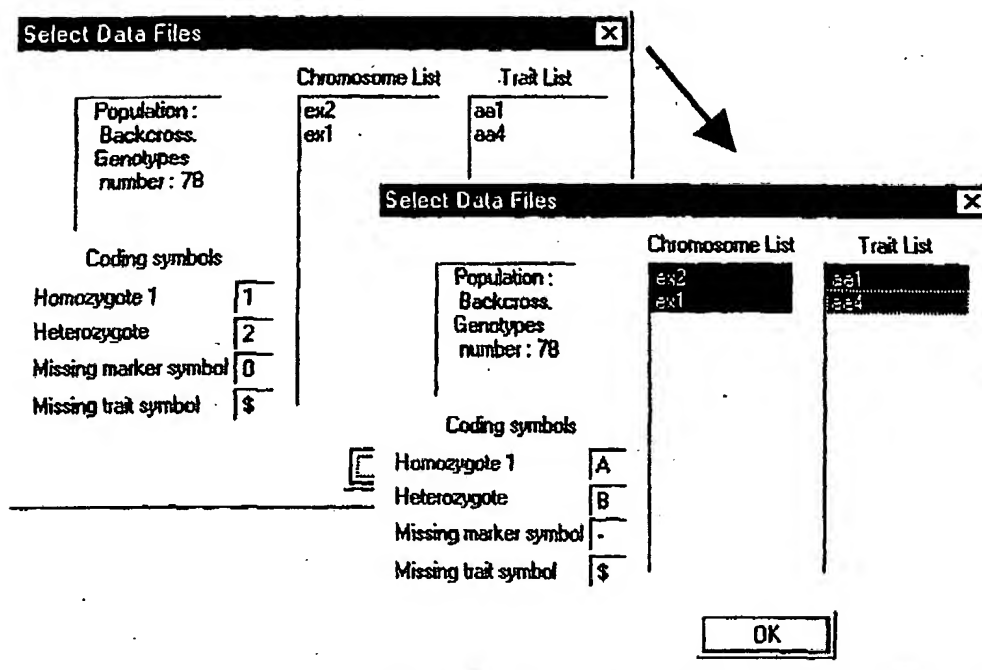
a single environment or multiple-environments.

14. A computer-readable storage system encoded with processing instructions for implementing a method for mapping qualitative trait loci in a mixed defined plant or animal population, said processing instructions for directing a computer to perform the steps of:

(a) selecting one or more traits across said population, thereby providing at I.A method for mapping quantitative trait Loci in a plant or animal population, the method comprising: east one quantified population phenotype;

(b) . identifying at least one genetic marker associated with the distribution of the trait, and

(d) (c) selecting the genotypes and markers defining one or more specific traits according to predetermined parameters.

15. The system of claim 14, wherein the plant or animal population includes:

(a) species with various reproductive systems including selfing and outbreeding;

(b) species with various types of life histories including annuals, perennials, diploids, polyploids or trees; or

(e) species of agricultural, industrial and ecological importance.

16. The system of claim 14, wherein the traits include:

yield, components and , quality;

resistance to abiotic and biotic stresses;

anatomical, physiological and biochemical characteristics;

quantitative expression patterns of different genes; and

traits with quantitative threshold manifestations.

17. The system of claim 14, wherein the mapping of the quantitative trait Loci

is carried out in one or more environments.

18. The system of claim 14, wherein the mapping of the quantitative trait Loci

is carried out during one or more stages of development of the population.

19. The system of claim 14, wherein the mapping of the quantitative trait Loci

is carried out in one or more families of the population.

20. The system of claim 14, wherein the mapping population includes

different population structures including backcross, F2 or F3.

21. The method of claim 1, wherein the mapping population includes

phenotypic data, or marker data from one or more generations.

22. The system of claim 14, wherein the mapping method includes: the

multilocus analysis of correlated trait complexes; the allowance for

variance and covariance effects; the discriminating between linkage and

pleiotropy; the testing for dominance and over dominance; and the

detection of epistasis.

23. The system of claim 14, wherein the method includes: the marker and

interval analysis, the maximum likelihood method, the method of moments

or the method of regression analysis.

24. The system of claim 14, wherein the association of the trait and the

genetic markers is determined by applying a statistical model or tools for

testing of significance and evaluation of precision.

25. The system of claim 14, wherein the tools for testing of significance and

evaluation of precision include: the permutation tests for the lod values

and user defined criteria; the bootstrap analysis ; the Monte-Carlo

simulations; or comparisions between different method and versions of the

same method.

26.   The system of claim 14, wherein the pre-determined parameters

include at least one of a group consisting of:

a single or multiple trait,

a single or linked QTL per chromosome, with or without fitting cofactors from

other chromosomes, and

a single environment or multiple-environments.

Multi QTL (unregistered version)

Data   Project   View   Model   Report   Help   Exit

FIGURE 1

Multi QTL

Data   Project   View   Model   Report   Help   Exit

Import   ►
Define Population Set   from ASCII file

FIGURE 2

Population Set Define   ☒

Population   Selective Genotyping

Backcross ▼   ☐

Total   Selected
Number   Objects

78   78

Enter Name of   qaz
Population Set

OK

FIGURE 3

FIGURE 4

SUBSTITUTE SHEET (RULE 26)

Input Data Report   ☒

Repair selected data file. Click 'Save' button to save your changes. Click 'Next' button to move to the next data file with errors. Click 'Finish' button to exit from errors repairing.

Commands

⦿ Edit Cell	○ Insert Cell
○ Row Split	○ Delete Cel

File Name

ex1_error

	1	2	3	4	5	▲
aa1	37.4	28.0	1	2	3	
aa4	28.9	25.0	26.0	25.6	24.5	▼

◄ |

Print   Save   Finish

FIGURE 5

Input Data Report   ☒

MultiQTL has detected some problems in your input data. To repair these problems click the 'Next' button.
To ignore files with errors - click the 'Cancel' button.

File name	Problem
⊗ ex1_error	Marker value is wrong in marker 7, object 11
⊗ ex1_error	Marker value is wrong in marker 8, object 7
⊗ ex1_error	Marker value is wrong in marker 8, object 12
⊗ ex1_error	Marker value is wrong in marker 8, object 13
⊗ ex1_error	Objects number in the marker # 8 AB6 = 80 instead of 78

Print | Next | Cancel

FIGURE 6

Input Data Report   ☒

Repair selected data file. Click 'Save' button to save your changes. Click 'Next' button to move to the
next data file with errors. Click 'Finish' button to exit from errors repairing.

Commands
● Edit Cell   ○ Insert Cell   File Name
○ Row Split   ○ Delete Cell   ex1_error

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
K94	A	A	A	A	A	A	B	A	B	B	B	B.	B	A	,
AR2	A	A	B	A	A	A	B	A	B	B	B	B	B	A	I
Ub1	A	A	B	-	B	A	B	-	-	-	-	B	-	A	:
A58	A	B	B	A	B	B	B	A	B	B	-	A	B	A	I
k1A	A	B	B	A	B	B	B	-	B	B	A	B	A	B	,
Q4A	A	B	B	A	B	B	B	A	B	B	A	B	A	B	,
A09	A	A	A	A	A	A	B	A	B	A	q	B	B	A	,
AB6	A	A	A	A	A	A	c	A	B	A	A	c	c	B	▼

Print | Save | Next | Finish

FIGURE 7

FIGURE 8

5/31

FIGURE 9

FIGURE 10

SUBSTITUTE SHEET (RULE 26)

FIGURE 11

SUBSTITUTE SHEET (RULE 26)

FIGURE 12

FIGURE 13

FIGURE 14

Is there any QTL interaction? (Y/N)

Yes   No

FIGURE 15

Epistas

QTL1   QTL2   QTL Pair Number
3   6   2
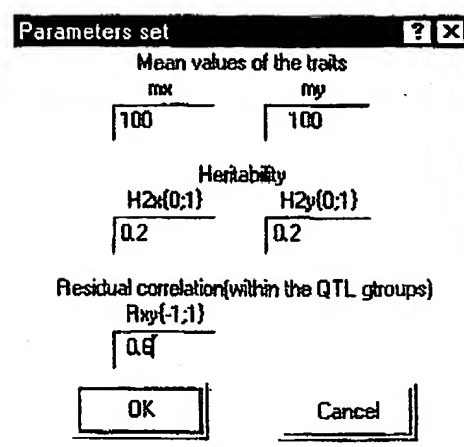
Epistas
0.30   0.10

QTL Pairs Scrolling
Add Pair   ◄ | ►   Delete Pair

FIGURE 16

Parameters set

Mean values of the traits
mx   my
100   100

Heritability
H2x(0;1)   H2y(0;1)
0.2   0.2

Residual correlation(within the QTL gtroups)
Rxy(-1;1)
0.6

OK   Cancel

9/31

FIGURE 17

FIGURE 18

FIGURE 19

FIGURE 20

SUBSTITUTE SHEET (RULE 26)

FIGURE 21

FIGURE 22

```
┌─────────────────────────────────────────────────────────────┐
│ Population :  Backcross                                    ☒ │
├─────────────────────────────────────────────────────────────┤
│ Model Parameters │ Extended Parameters │                     │
│                                                               │
│  Model Name      │d1│         Number of Traits  │One     ▼│  │
│                                                               │
│  Mapping Function Haldane  ▼│  Names of Traits   ┌trait1───┐  │
│                                                   │trait2   │  │
│  Number of QTL  │Single QTl ▼│                   └─────────┘  │
│                                                               │
│                                  Selective Genotyping         │
│  Initial Submodel Default ☑              ☐  │ On/Off │        │
│                                                               │
│                      ☐           Selected Trait ┌─────────┐   │
│                                                 └─────────┘   │
│         Models Scrolling                                      │
│         ◄    │    ►            Add Model ║  Delete Model ║     │
│                                                               │
├─────────────────────────────────────────────────────────────┤
│                      │   OK    │   Cancel  │                  │
└─────────────────────────────────────────────────────────────┘
```
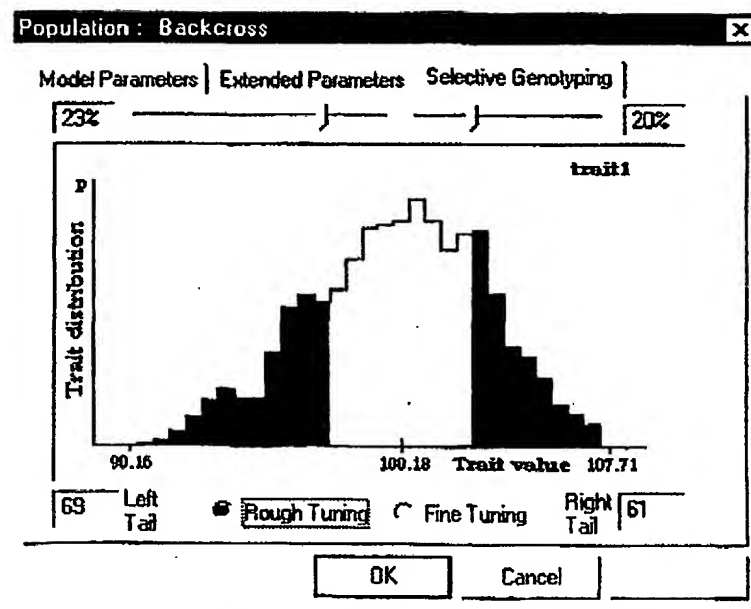
FIGURE 23

FIGURE 24

FIGURE 25

FIGURE 26

SUBSTITUTE SHEET (RULE 26)

FIGURE 27

```
┌─────────────────────────────────────┐
│ About model                      [X] │
│ ┌──────────────────────────────────┐ │
│ │ Model d2 contains :              │ │
│ │ 3 chromosome(s),  2 trait(s)     │ │
│ │ Objects number :   300           │ │
│ │ Mapping function : Haldane       │ │
│ │ Population :       Backcross      │ │
│ │ Traits analysis :  Two traits    │ │
│ │ QTL analysis :     Single QTL    │ │
│ │ Intervals analysis               │ │
│ │ Lost Markers Percent : 3         │ │
│ │ Without markers restoration      │ │
│ │                                  │ │
│ │                                  │ │
│ │                                  │ │
│ └──────────────────────────────────┘ │
│ ┌──────────────┐                      │
│ │About Simulation│                    │
│ └──────────────┘                      │
└─────────────────────────────────────┘
```
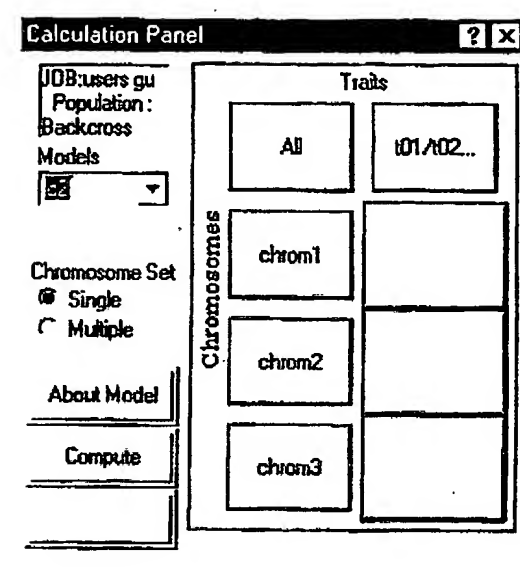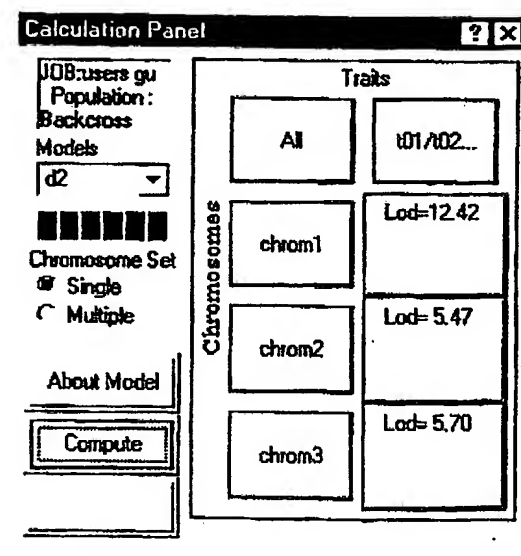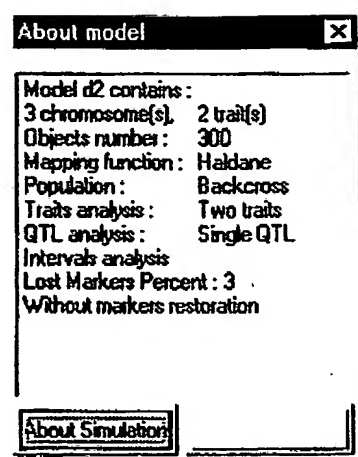
```
┌──────────────────────────┐
│ ⎺ Specified         [X]   │
│ Interval 4 (5.35cM)       │
│      Parameters           │
│ ┌────┬───────┬───────┐    │
│ │ E..│ trait1│ trait2│    │
│ ├────┼───────┼───────┤    │
│ │ d  │ 1.000 │ -1.000│    │
│ └────┴───────┴───────┘    │
└──────────────────────────┘
```

FIGURE 28

```
┌───────────────────────────────────────────────────────────────────┐
│ ⎺ Occurred                                                    [X]   │
│ Interval 4 (5.27cM)                                                 │
│     Parameters            Std. Dev.            Cor. Coef.           │
│ ┌────┬────────┬─────────┬────┬───────┬───────┬────┬───────┐        │
│ │ E..│ trait1 │ trait2  │ G..│ trait1│ trait2│ G..│ t1/t2 │        │
│ ├────┼────────┼─────────┼────┼───────┼───────┼────┼───────┤        │
│ │ m  │100.082 │100.098  │ Qq │ 3.186 │ 2.527 │ Qq │ 0.524 │        │
│ │ d  │  1.550 │ -0.523  │ qq │ 3.132 │ 2.354 │ qq │ 0.496 │        │
│ └────┴────────┴─────────┴────┴───────┴───────┴────┴───────┘        │
└───────────────────────────────────────────────────────────────────┘
```

FIGURE 29

16/31

FIGURE 30

Submodel   ☒

01trait1   02trait2
  Effect   Effect

  ○ effect1=0   ○ effect1=0
  ○ effect2=0   ○ effect2=0
  ○ epistasis=0   ○ epistasis=0

  Variance Effect   Variance Effect

  ◉ no variance effect   ◉ no variance effect

  Covariance Effect

  ◉ no covariance effect

  [OK   ]   Cancel

Submodel   ☒

  Effect
  ○ no effect(d=0)

  Variance Effect
  ◉ no variance effect

  [OK   ]   Cancel

FIGURE 31

FIGURE 32

Multi QTL

Data Project View

ASCII ⊕
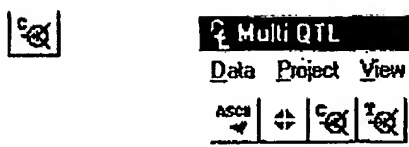
FIGURE 33

Chromosome information

Chromosomes [▼]   Apply | Close |

Number of markers :14

Marker number	a	1	2	3	4	5	6	7	8	9	10
Defined objects on marker		78	78	69	74	74	74	78	77	77	78
Interval length		24.3	33.9	140.2	97.3	42.4	145.9	1.3	2.7	2.7	
Defined objects on interval		78	69	66	70	70	74	77	76	77	

Change markers
- Select marker to delete
-
- Fill missing

Segregation

1 K94	2 AR2	3 Ub1	4 A58
1.0:0.9 $x^2 = 0.5$	1.0:0.9 $x^2 = 1.1$	0.9:1.0 $x^2 = 0.8$	1.0:0.9 $x^2 = 0.5$

FIGURE 34

Changed Chromosome Report

Number of markers :13

Deleted marker 3 Ub1

1	2	4	5	6	7	8	9	10	11
78	78	74	74	74	78	77	77	78	78
24.3	125.6	97.3	42.4	145.9	1.3	2.7	2.7	6.9	
78	74	70	70	74	77	76	77	78	

18/31

FIGURE 35

Status of the new chromosome   ☒

    ○ Save a new chromosome and add it to job file

    ◉ Add a new chromosome to job file without saving

    ○ Replace a starting chromosome by new one and save it

    ○ Exit without saving

    | OK | New chromosome name | ex1_a |

FIGURE 36

Fill missing markers parameters   ☒

0 0 0 0 1 2 1 0 0 0 0 0 2 1 3 1 0 0 0 0 0

External interval   Internal interval   External interval

Units

○ MU   ◉ CM

Max. Length of External   Max. Length of Internal
Interval Uncertainty   Interval Uncertainty

| ▓▓▓ | | 0.00 | | OK | Cancel

FIGURE 37
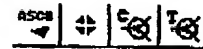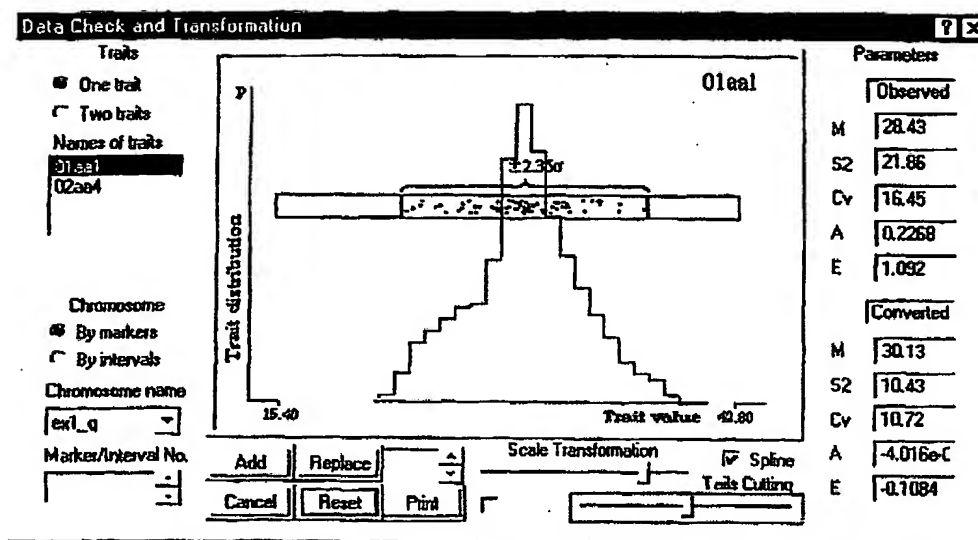
FIGURE 38

No.	Trait value
39	20.763
44	15.400
51	18.200

FIGURE 39

20/31
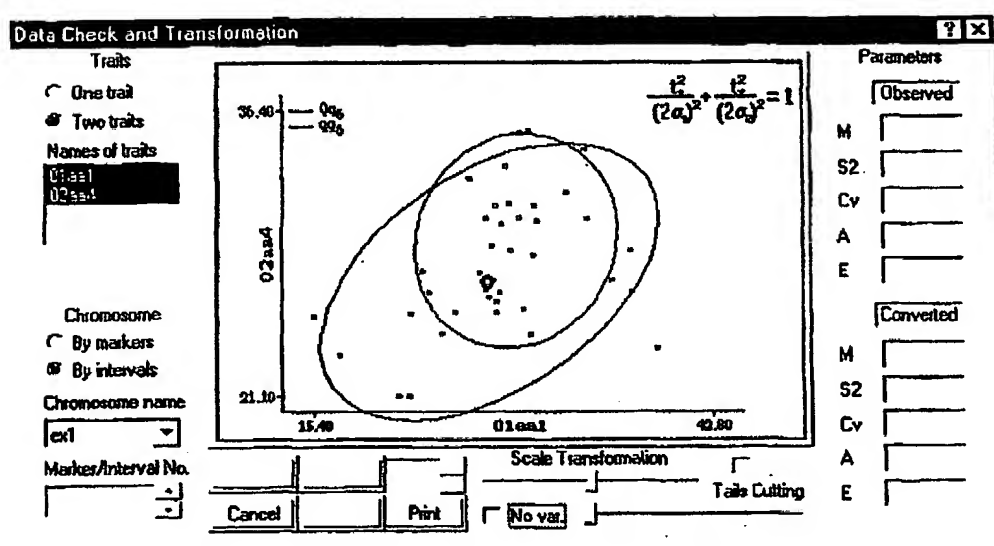
FIGURE 40

FIGURE 41
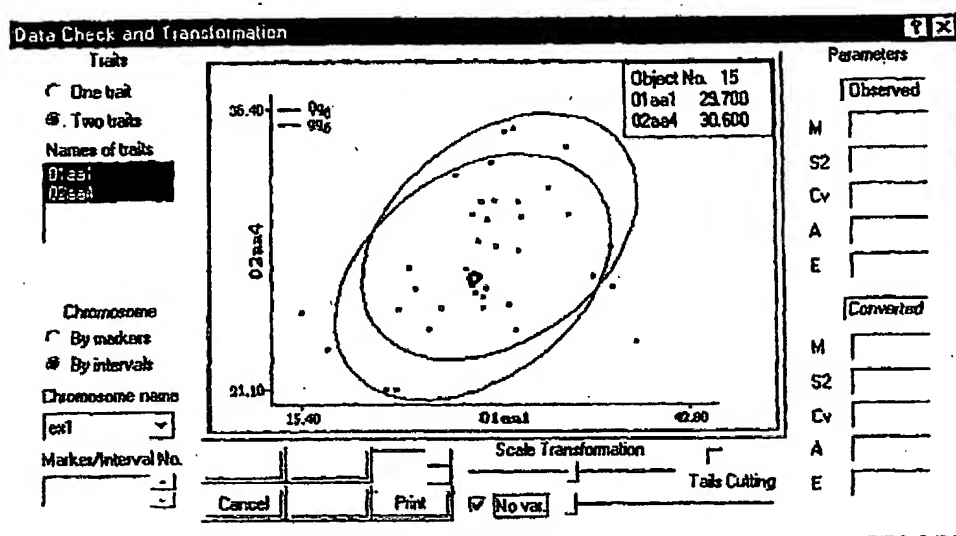
FIGURE 42
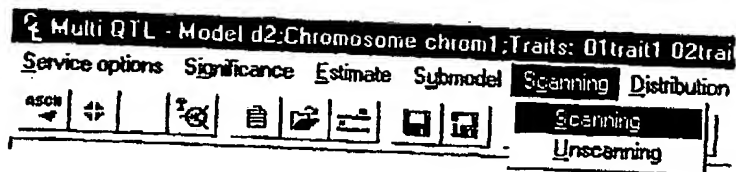
FIGURE 43
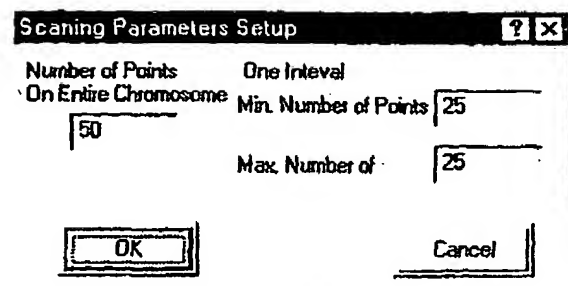
FIGURE 44
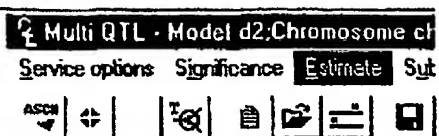
FIGURE 45
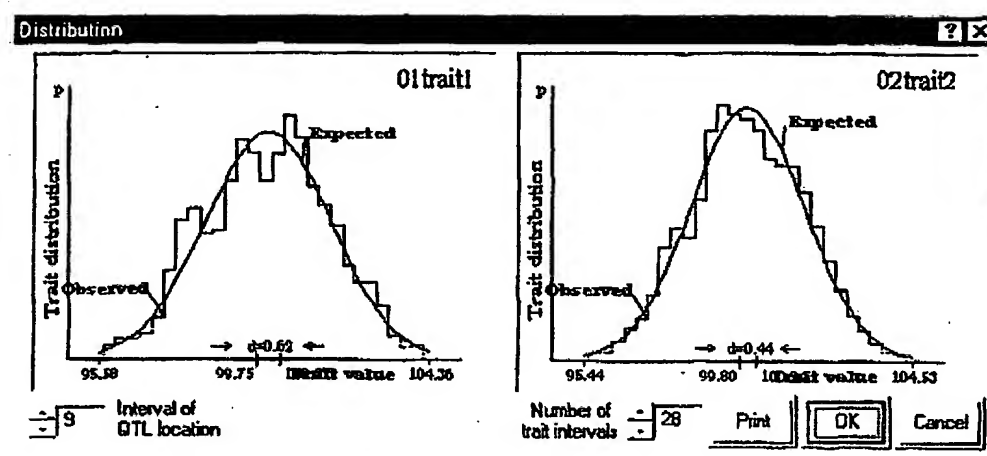
FIGURE 46
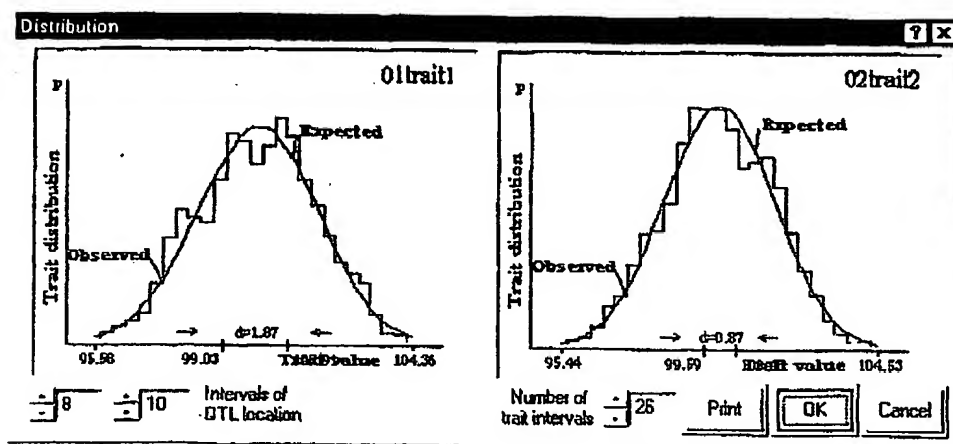
FIGURE 47

FIGURE 48

FIGURE 49

FIGURE 50

25/31

FIGURE 51

Bootstrap test

	L(cM)	tr 1 (01trait1) Parameters aver.(m)	eff.(d)	Std. Dev.	tr 2 (02trait2) Parameters aver.(m)	eff.(d)	Std. Dev.
Mean value	52.077	100.052	-0.398	1.612	100.010	-0.222	1.542
Stand.dev.	43.846	0.189	1.251	0.123	0.177	1.077	0.121

Number of samples 1000 / 1000 ▆▆▆▆▆▆▆▆▆▆▆▆▆ Filter (%) 100 Reset Start Close Print

FIGURE 52

Bootstrap test

	L1(cM)	L2(cM)	tr 1 (01trait1) Parameters aver(m)	eff1(d1)	eff2(d2)	epist.(e)	Std. Dev.
Mean value	23.494	61.987	100.158	0.870	-1.041	-0.152	1.536
Stand.dev.	29.156	15.213	0.365	1.475	1.368	0.430	0.127

Number of samples 1000 / 200 ▆▆▆▆ Filter (%) 100 Reset Start Close Print

FIGURE 53

FIGURE 54

FIGURE 55

FIGURE 56

FIGURE 57
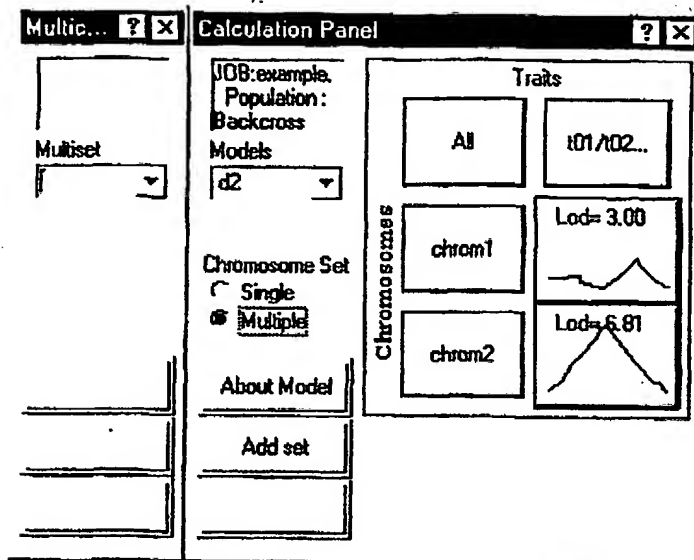
FIGURE 58
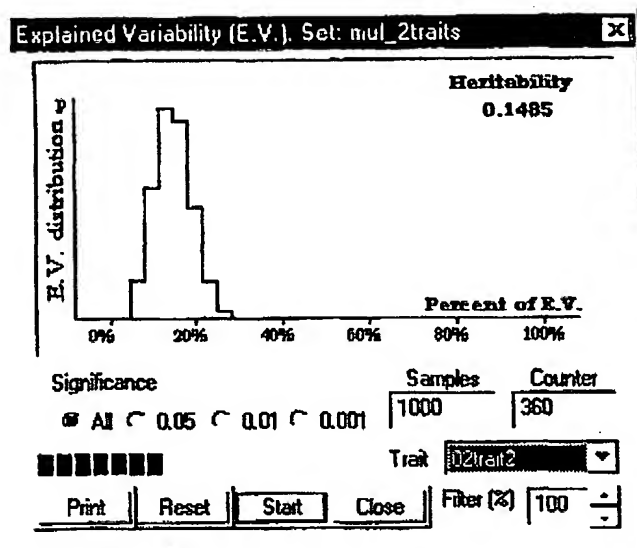
FIGURE 59

FIGURE 60

	L(cM)	tr 1 (01trait1)			tr 2 (02trait2)		
		Effect		Variance	Effect		Variance
		aver.[m]	eff.[d]		aver.[m]	eff.[d]	
Mean value	55.252	100.007	1.029	1.627	99.999	1.018	1.629
Stand.dev.	15.876	0.186	0.406	0.169	0.186	0.397	0.169

FIGURE 61

	L1(cM)	L2(cM)	tr 1 (01trait1)				Variance
			Effect				
			aver(m)	eff1(d1)	eff2(d2)	epist.(e)	
Mean value	41.033	67.260	99.991	0.482	0.522	-0.007	1.614
Stand.dev.	39.779	35.044	0.458	1.343	1.340	0.526	0.170

FIGURE 62

Multi QTL - Model d4;0

Service options | Significance

Print...	Ctrl+P
Print Preview	
Print Setup...	
Save job	
Save job as...	
Make Report	
Exit	

FIGURE 63

Print file ☒

● Print

○ Save to PCX file

○ Change Font

[OK]

FIGURE 64

:Chromosome chrom2;Traits: 01trait1 ☒

MultiQTL Ltd.
Contact korol@multiqtl.com
or Tel: 972-4-842-3739

[Save&Exit]

31/31